



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:32 PM GMT

PDB ID : 4LW5  
Title : Crystal structure of all-trans green fluorescent protein  
Authors : Rosenman, D.J.; Huang, Y.-M.; Xia, K.; Vanroey, P.; Colon, W.; Bystroff, C.  
Deposited on : 2013-07-26  
Resolution : 2.55 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

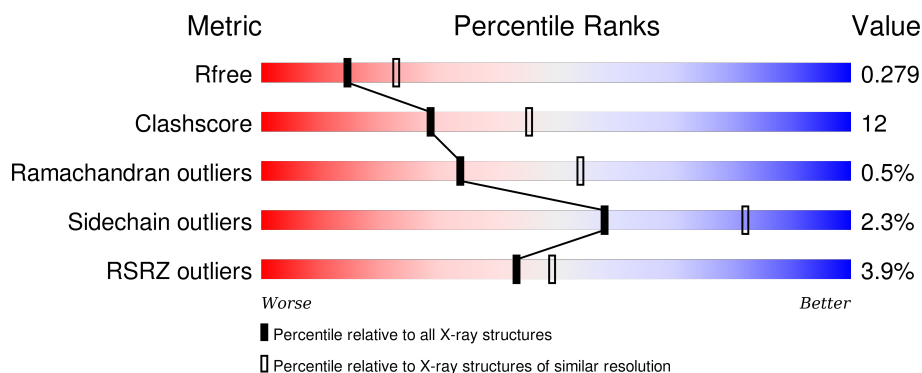
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	4549 (2.58-2.50)
Clashscore	102246	5292 (2.58-2.50)
Ramachandran outliers	100387	5194 (2.58-2.50)
Sidechain outliers	100360	5196 (2.58-2.50)
RSRZ outliers	91569	4561 (2.58-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	239	<div> <div>2%</div> <div> <div></div> <div>73%</div> <div>20%</div> <div>• 5%</div> </div> </div>
1	B	239	<div> <div>6%</div> <div> <div></div> <div>70%</div> <div>25%</div> <div>5%</div> </div> </div>
1	C	239	<div> <div>3%</div> <div> <div></div> <div>77%</div> <div>15%</div> <div>• 5%</div> </div> </div>
1	D	239	<div> <div>3%</div> <div> <div></div> <div>72%</div> <div>21%</div> <div>• 6%</div> </div> </div>
1	E	239	<div> <div>4%</div> <div> <div></div> <div>72%</div> <div>22%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9411 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total	C	N	O	S	0	0	0
			1797	1139	309	345	4			
1	B	227	Total	C	N	O	S	0	0	0
			1807	1145	312	346	4			
1	C	226	Total	C	N	O	S	0	0	0
			1794	1137	309	344	4			
1	D	225	Total	C	N	O	S	0	0	0
			1788	1134	308	342	4			
1	E	226	Total	C	N	O	S	0	0	0
			1797	1139	309	345	4			

There are 145 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP P42212
A	1	ALA	-	EXPRESSION TAG	UNP P42212
A	30	ARG	SER	ENGINEERED MUTATION	UNP P42212
A	39	ILE	TYR	ENGINEERED MUTATION	UNP P42212
A	64	LEU	PHE	ENGINEERED MUTATION	UNP P42212
A	66	CRO	SER	CHROMOPHORE	UNP P42212
A	66	CRO	TYR	CHROMOPHORE	UNP P42212
A	66	CRO	GLY	CHROMOPHORE	UNP P42212
A	80	ARG	GLN	ENGINEERED MUTATION	UNP P42212
A	88	ILE	-	INSERTION	UNP P42212
A	89	SER	-	INSERTION	UNP P42212
A	90	ASN	MET	ENGINEERED MUTATION	UNP P42212
A	91	GLY	PRO	ENGINEERED MUTATION	UNP P42212
A	92	ASP	GLU	ENGINEERED MUTATION	UNP P42212
A	94	PHE	TYR	ENGINEERED MUTATION	UNP P42212
A	95	ILE	VAL	ENGINEERED MUTATION	UNP P42212
A	96	ASN	GLN	ENGINEERED MUTATION	UNP P42212
A	101	SER	PHE	ENGINEERED MUTATION	UNP P42212
A	107	LYS	ASN	ENGINEERED MUTATION	UNP P42212

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	113	VAL	GLU	ENGINEERED MUTATION	UNP P42212
A	130	THR	ILE	ENGINEERED MUTATION	UNP P42212
A	147	PHE	TYR	ENGINEERED MUTATION	UNP P42212
A	155	THR	MET	ENGINEERED MUTATION	UNP P42212
A	165	ALA	VAL	ENGINEERED MUTATION	UNP P42212
A	168	THR	LYS	ENGINEERED MUTATION	UNP P42212
A	169	VAL	ILE	ENGINEERED MUTATION	UNP P42212
A	173	VAL	ILE	ENGINEERED MUTATION	UNP P42212
A	207	THR	SER	ENGINEERED MUTATION	UNP P42212
A	208	VAL	ALA	ENGINEERED MUTATION	UNP P42212
B	0	MET	-	EXPRESSION TAG	UNP P42212
B	1	ALA	-	EXPRESSION TAG	UNP P42212
B	30	ARG	SER	ENGINEERED MUTATION	UNP P42212
B	39	ILE	TYR	ENGINEERED MUTATION	UNP P42212
B	64	LEU	PHE	ENGINEERED MUTATION	UNP P42212
B	66	CRO	SER	CHROMOPHORE	UNP P42212
B	66	CRO	TYR	CHROMOPHORE	UNP P42212
B	66	CRO	GLY	CHROMOPHORE	UNP P42212
B	80	ARG	GLN	ENGINEERED MUTATION	UNP P42212
B	88	ILE	-	INSERTION	UNP P42212
B	89	SER	-	INSERTION	UNP P42212
B	90	ASN	MET	ENGINEERED MUTATION	UNP P42212
B	91	GLY	PRO	ENGINEERED MUTATION	UNP P42212
B	92	ASP	GLU	ENGINEERED MUTATION	UNP P42212
B	94	PHE	TYR	ENGINEERED MUTATION	UNP P42212
B	95	ILE	VAL	ENGINEERED MUTATION	UNP P42212
B	96	ASN	GLN	ENGINEERED MUTATION	UNP P42212
B	101	SER	PHE	ENGINEERED MUTATION	UNP P42212
B	107	LYS	ASN	ENGINEERED MUTATION	UNP P42212
B	113	VAL	GLU	ENGINEERED MUTATION	UNP P42212
B	130	THR	ILE	ENGINEERED MUTATION	UNP P42212
B	147	PHE	TYR	ENGINEERED MUTATION	UNP P42212
B	155	THR	MET	ENGINEERED MUTATION	UNP P42212
B	165	ALA	VAL	ENGINEERED MUTATION	UNP P42212
B	168	THR	LYS	ENGINEERED MUTATION	UNP P42212
B	169	VAL	ILE	ENGINEERED MUTATION	UNP P42212
B	173	VAL	ILE	ENGINEERED MUTATION	UNP P42212
B	207	THR	SER	ENGINEERED MUTATION	UNP P42212
B	208	VAL	ALA	ENGINEERED MUTATION	UNP P42212
C	0	MET	-	EXPRESSION TAG	UNP P42212
C	1	ALA	-	EXPRESSION TAG	UNP P42212
C	30	ARG	SER	ENGINEERED MUTATION	UNP P42212

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	39	ILE	TYR	ENGINEERED MUTATION	UNP P42212
C	64	LEU	PHE	ENGINEERED MUTATION	UNP P42212
C	66	CRO	SER	CHROMOPHORE	UNP P42212
C	66	CRO	TYR	CHROMOPHORE	UNP P42212
C	66	CRO	GLY	CHROMOPHORE	UNP P42212
C	80	ARG	GLN	ENGINEERED MUTATION	UNP P42212
C	88	ILE	-	INSERTION	UNP P42212
C	89	SER	-	INSERTION	UNP P42212
C	90	ASN	MET	ENGINEERED MUTATION	UNP P42212
C	91	GLY	PRO	ENGINEERED MUTATION	UNP P42212
C	92	ASP	GLU	ENGINEERED MUTATION	UNP P42212
C	94	PHE	TYR	ENGINEERED MUTATION	UNP P42212
C	95	ILE	VAL	ENGINEERED MUTATION	UNP P42212
C	96	ASN	GLN	ENGINEERED MUTATION	UNP P42212
C	101	SER	PHE	ENGINEERED MUTATION	UNP P42212
C	107	LYS	ASN	ENGINEERED MUTATION	UNP P42212
C	113	VAL	GLU	ENGINEERED MUTATION	UNP P42212
C	130	THR	ILE	ENGINEERED MUTATION	UNP P42212
C	147	PHE	TYR	ENGINEERED MUTATION	UNP P42212
C	155	THR	MET	ENGINEERED MUTATION	UNP P42212
C	165	ALA	VAL	ENGINEERED MUTATION	UNP P42212
C	168	THR	LYS	ENGINEERED MUTATION	UNP P42212
C	169	VAL	ILE	ENGINEERED MUTATION	UNP P42212
C	173	VAL	ILE	ENGINEERED MUTATION	UNP P42212
C	207	THR	SER	ENGINEERED MUTATION	UNP P42212
C	208	VAL	ALA	ENGINEERED MUTATION	UNP P42212
D	0	MET	-	EXPRESSION TAG	UNP P42212
D	1	ALA	-	EXPRESSION TAG	UNP P42212
D	30	ARG	SER	ENGINEERED MUTATION	UNP P42212
D	39	ILE	TYR	ENGINEERED MUTATION	UNP P42212
D	64	LEU	PHE	ENGINEERED MUTATION	UNP P42212
D	66	CRO	SER	CHROMOPHORE	UNP P42212
D	66	CRO	TYR	CHROMOPHORE	UNP P42212
D	66	CRO	GLY	CHROMOPHORE	UNP P42212
D	80	ARG	GLN	ENGINEERED MUTATION	UNP P42212
D	88	ILE	-	INSERTION	UNP P42212
D	89	SER	-	INSERTION	UNP P42212
D	90	ASN	MET	ENGINEERED MUTATION	UNP P42212
D	91	GLY	PRO	ENGINEERED MUTATION	UNP P42212
D	92	ASP	GLU	ENGINEERED MUTATION	UNP P42212
D	94	PHE	TYR	ENGINEERED MUTATION	UNP P42212
D	95	ILE	VAL	ENGINEERED MUTATION	UNP P42212

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	96	ASN	GLN	ENGINEERED MUTATION	UNP P42212
D	101	SER	PHE	ENGINEERED MUTATION	UNP P42212
D	107	LYS	ASN	ENGINEERED MUTATION	UNP P42212
D	113	VAL	GLU	ENGINEERED MUTATION	UNP P42212
D	130	THR	ILE	ENGINEERED MUTATION	UNP P42212
D	147	PHE	TYR	ENGINEERED MUTATION	UNP P42212
D	155	THR	MET	ENGINEERED MUTATION	UNP P42212
D	165	ALA	VAL	ENGINEERED MUTATION	UNP P42212
D	168	THR	LYS	ENGINEERED MUTATION	UNP P42212
D	169	VAL	ILE	ENGINEERED MUTATION	UNP P42212
D	173	VAL	ILE	ENGINEERED MUTATION	UNP P42212
D	207	THR	SER	ENGINEERED MUTATION	UNP P42212
D	208	VAL	ALA	ENGINEERED MUTATION	UNP P42212
E	0	MET	-	EXPRESSION TAG	UNP P42212
E	1	ALA	-	EXPRESSION TAG	UNP P42212
E	30	ARG	SER	ENGINEERED MUTATION	UNP P42212
E	39	ILE	TYR	ENGINEERED MUTATION	UNP P42212
E	64	LEU	PHE	ENGINEERED MUTATION	UNP P42212
E	66	CRO	SER	CHROMOPHORE	UNP P42212
E	66	CRO	TYR	CHROMOPHORE	UNP P42212
E	66	CRO	GLY	CHROMOPHORE	UNP P42212
E	80	ARG	GLN	ENGINEERED MUTATION	UNP P42212
E	88	ILE	-	INSERTION	UNP P42212
E	89	SER	-	INSERTION	UNP P42212
E	90	ASN	MET	ENGINEERED MUTATION	UNP P42212
E	91	GLY	PRO	ENGINEERED MUTATION	UNP P42212
E	92	ASP	GLU	ENGINEERED MUTATION	UNP P42212
E	94	PHE	TYR	ENGINEERED MUTATION	UNP P42212
E	95	ILE	VAL	ENGINEERED MUTATION	UNP P42212
E	96	ASN	GLN	ENGINEERED MUTATION	UNP P42212
E	101	SER	PHE	ENGINEERED MUTATION	UNP P42212
E	107	LYS	ASN	ENGINEERED MUTATION	UNP P42212
E	113	VAL	GLU	ENGINEERED MUTATION	UNP P42212
E	130	THR	ILE	ENGINEERED MUTATION	UNP P42212
E	147	PHE	TYR	ENGINEERED MUTATION	UNP P42212
E	155	THR	MET	ENGINEERED MUTATION	UNP P42212
E	165	ALA	VAL	ENGINEERED MUTATION	UNP P42212
E	168	THR	LYS	ENGINEERED MUTATION	UNP P42212
E	169	VAL	ILE	ENGINEERED MUTATION	UNP P42212
E	173	VAL	ILE	ENGINEERED MUTATION	UNP P42212
E	207	THR	SER	ENGINEERED MUTATION	UNP P42212
E	208	VAL	ALA	ENGINEERED MUTATION	UNP P42212

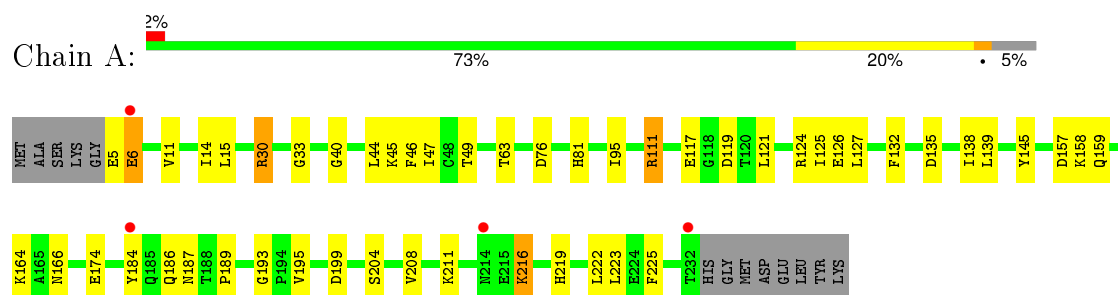
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	116	Total 116	O 116	0	0
2	B	102	Total 102	O 102	0	0
2	C	84	Total 84	O 84	0	0
2	D	64	Total 64	O 64	0	0
2	E	62	Total 62	O 62	0	0

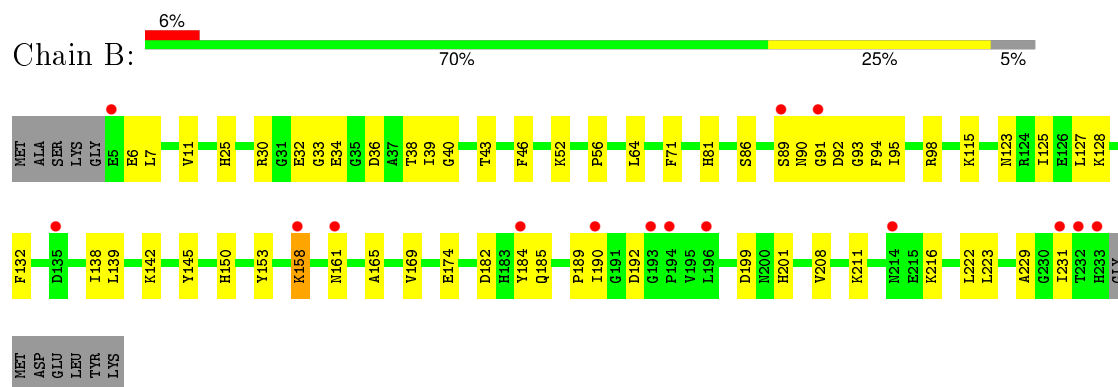
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

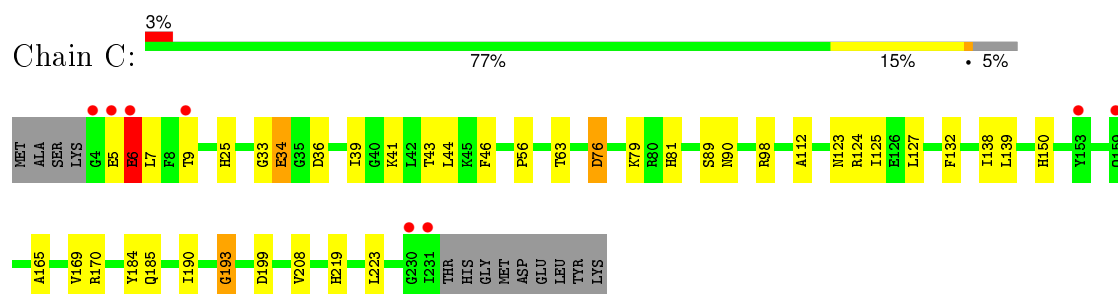
- Molecule 1: Green fluorescent protein



- Molecule 1: Green fluorescent protein



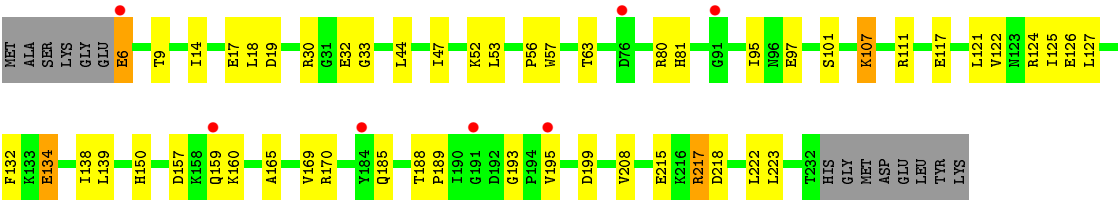
- Molecule 1: Green fluorescent protein



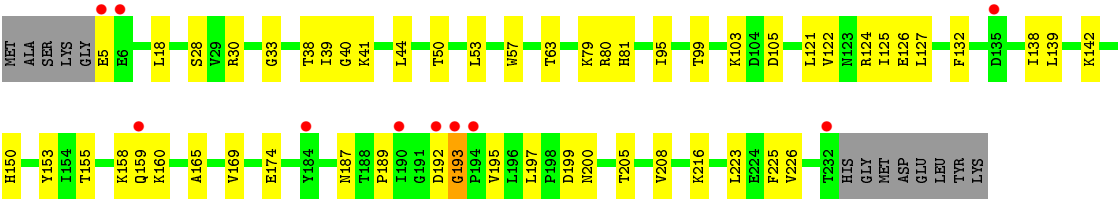
- Molecule 1: Green fluorescent protein







● Molecule 1: Green fluorescent protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.56Å 110.94Å 114.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.94 – 2.55 19.94 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.94-2.55) 99.8 (19.94-2.55)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.49 (at 2.56Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.218 , 0.281 0.218 , 0.279	Depositor DCC
$R_{free}$ test set	1999 reflections (4.97%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.4	Xtriage
Anisotropy	0.406	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 39.0	EDS
Estimated twinning fraction	0.005 for -h,l,k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 40237 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	9411	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CRO

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.41	0/1812	0.67	0/2451
1	B	0.41	0/1823	0.65	0/2466
1	C	0.41	0/1809	0.66	0/2446
1	D	0.42	0/1803	0.65	0/2439
1	E	0.41	0/1812	0.67	1/2451 (0.0%)
All	All	0.41	0/9059	0.66	1/12253 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	99	THR	N-CA-C	-5.08	97.29	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1797	0	1754	49	0
1	B	1807	0	1763	53	0
1	C	1794	0	1752	34	0
1	D	1788	0	1750	40	0
1	E	1797	0	1756	39	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	116	0	0	6	0
2	B	102	0	0	2	0
2	C	84	0	0	5	0
2	D	64	0	0	3	0
2	E	62	0	0	1	0
All	All	9411	0	8775	205	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 12.

The worst 5 of 205 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:184:TYR:CE2	1:B:190:ILE:HB	1.88	1.07
1:B:158:LYS:HD3	1:B:158:LYS:H	1.33	0.93
1:B:182:ASP:OD2	1:B:184:TYR:HE1	1.62	0.81
1:B:182:ASP:OD2	1:B:184:TYR:CE1	2.35	0.80
1:D:215:GLU:OE2	1:D:217:ARG:HG2	1.82	0.79

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	221/239 (92%)	213 (96%)	7 (3%)	1 (0%)	34	54
1	B	222/239 (93%)	211 (95%)	10 (4%)	1 (0%)	34	54
1	C	221/239 (92%)	210 (95%)	9 (4%)	2 (1%)	21	36
1	D	220/239 (92%)	212 (96%)	7 (3%)	1 (0%)	34	54
1	E	221/239 (92%)	212 (96%)	8 (4%)	1 (0%)	34	54
All	All	1105/1195 (92%)	1058 (96%)	41 (4%)	6 (0%)	34	54

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	6	GLU
1	C	193	GLY
1	E	193	GLY
1	A	193	GLY
1	B	91	GLY

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	197/207 (95%)	189 (96%)	8 (4%)	37	61
1	B	198/207 (96%)	193 (98%)	5 (2%)	55	80
1	C	196/207 (95%)	193 (98%)	3 (2%)	72	90
1	D	196/207 (95%)	190 (97%)	6 (3%)	47	73
1	E	197/207 (95%)	196 (100%)	1 (0%)	92	98
All	All	984/1035 (95%)	961 (98%)	23 (2%)	58	82

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	52	LYS
1	B	192	ASP
1	D	217	ARG
1	B	158	LYS
1	C	6	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	161	ASN
1	E	159	GLN
1	C	81	HIS
1	B	25	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	25	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

5 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	CRO	A	66	1	23,23,24	7.09	14 (60%)	29,32,34	3.43	15 (51%)
1	CRO	B	66	1	23,23,24	4.80	13 (56%)	29,32,34	1.99	9 (31%)
1	CRO	C	66	1	23,23,24	6.08	18 (78%)	29,32,34	3.20	16 (55%)
1	CRO	D	66	1	23,23,24	6.24	16 (69%)	29,32,34	4.06	17 (58%)
1	CRO	E	66	1	23,23,24	5.29	13 (56%)	29,32,34	2.62	15 (51%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CRO	A	66	1	-	0/12/31/32	0/2/2/2
1	CRO	B	66	1	-	0/12/31/32	0/2/2/2
1	CRO	C	66	1	-	0/12/31/32	0/2/2/2
1	CRO	D	66	1	-	0/12/31/32	0/2/2/2
1	CRO	E	66	1	-	0/12/31/32	0/2/2/2

The worst 5 of 74 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	66	CRO	OH-CZ	-8.28	1.17	1.37
1	D	66	CRO	CA1-C1	-7.81	1.39	1.51
1	A	66	CRO	CG2-CB2	-7.07	1.32	1.46
1	A	66	CRO	CA3-N3	-6.16	1.37	1.47
1	B	66	CRO	CA1-C1	-5.86	1.42	1.51

The worst 5 of 72 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	66	CRO	N3-C1-N2	-12.53	101.88	111.56
1	C	66	CRO	CA2-C2-N3	-9.00	98.89	103.40
1	A	66	CRO	CA1-C1-N2	-8.73	111.84	123.83
1	A	66	CRO	O2-C2-CA2	-8.43	126.39	130.95
1	D	66	CRO	O2-C2-CA2	-6.59	127.39	130.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	225/239 (94%)	-0.04	4 (1%) 71 76	17, 31, 48, 63	0
1	B	226/239 (94%)	0.21	15 (6%) 22 25	18, 33, 60, 84	0
1	C	225/239 (94%)	0.03	8 (3%) 46 53	17, 33, 53, 79	0
1	D	224/239 (93%)	0.13	7 (3%) 52 58	18, 35, 58, 67	0
1	E	225/239 (94%)	0.16	10 (4%) 38 44	19, 35, 63, 76	0
All	All	1125/1195 (94%)	0.10	44 (3%) 43 49	17, 33, 58, 84	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	184	TYR	4.7
1	C	4	GLY	4.4
1	B	196	LEU	4.3
1	A	184	TYR	4.3
1	A	232	THR	4.1

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	CRO	D	66	22/23	0.94	0.15	-	19,23,27,31	0
1	CRO	B	66	22/23	0.96	0.12	-	18,21,24,25	0
1	CRO	E	66	22/23	0.94	0.14	-	20,24,28,28	0
1	CRO	C	66	22/23	0.95	0.15	-	19,23,25,26	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	CRO	A	66	22/23	0.96	0.13	-	19,22,26,27	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.